Ground Forecasting in Mechanized Tunneling

Saadeldin Mostafa[1], Rita L. Sousa[1](✉), Herbert H. Einstein[2], and Beatriz G. Klink[2]

[1] Stevens Institute of Technology, Hoboken, NJ 07030, USA
rsousa@stevens.edu
[2] Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract. The construction of TBM tunnels is associated with high uncertainty due to the unknown ground conditions surrounding the TBM. Recently, there have been several attempts to make use of the large amount of TBM data recorded during construction to predict the ground conditions and automate the tunneling process. This study presents an implementation of supervised learning models to the Porto metro dataset (Sousa and Einstein 2012) and showcases an alternative method of predicting the ground class. The results of several machine learning (ML) models are reported and compared to each other. These ML models use the same algorithm but with different sets of input features (i.e., TBM parameters) to investigate the effect of different TBM parameters on predicting the geology of the tunnel. The results show that the learned model achieved high accuracy when predicting ground classes. Also, it indicates that input feature selection process is a crucial step to build a robust model since it eliminates ambiguous data thus increasing modeling accuracy while reducing training time. Moreover, the confusion matrices of different models showed that the rock ground class had higher scores and consistency under different sets of TBM features. This suggests that other ground classes need more refinement to attain a better model performance.

Keywords: Ground Prediction · TBM · Machine Learning

1 Introduction

The past few years have witnessed an increase in the number of tunnels as well as rapid development in the field of construction technology, particularly mechanized tunnel construction. Construction with tunnel boring machines (TBMs) is safer and faster when compared to conventional tunneling methods. Nevertheless, TBM control is a complex process due to the uncertainties associated with unforeseen geology ahead of the TBM (Sousa and Einstein 2021). The human operator's goal is to keep the operational parameters within predefined thresholds to achieve an optimal and safe operation. However, these thresholds are set during the design phase based on the small amount of geologic data collected during site investigation, which are often different than actual geology. The human operators rely on their experience and judgement to infer the ground type based on the readings of the TBM sensors and adjust the operational parameters to meet the sensed field conditions. The uncertainties associated with the TBM control translates

S. Javankhoshdel and Y. Abolfazlzadeh (Eds.): TMIC 2022, AHE 13, pp. 240–252, 2023.
https://doi.org/10.2991/978-94-6463-104-3_21

into risks that may cause an increase in time, and costs of the projects as well as accidents (Sousa and Einstein 2021).

In the past decades, researchers have been developing geology prediction models to assist the operator and to minimize risks during construction. These models can be broadly classified into probabilistic, and data driven or machine learning (ML) models. The probabilistic models include Markov Chain models (Chan 1981) and Dynamic Bayesian models (Sousa and Einstein 2012), among others. ML models can be further classified according to how they are learned: supervised, unsupervised, or semi-supervised learning models (Bishop 2006). Among the most popular ML algorithms are Artificial Neural Networks (ANN), Random Forests (RN), K-nearest neighbor (KNN) and Support Vector Machines (SVM) (e.g., Zhang et al. 2020; Shi et al. 2019; Zhao et al. 2019; Hou et al. 2021; Liu et al. 2020; Wu et al. 2021; Yin et al. 2022).

The existing models, both probabilistic and data driven, have several shortcomings that often do not make them suitable for tunneling automation. One such limitation is that existing models do not consider all the parameters recorded by the TBM, choosing the ones to include (i.e., the input features) in the models sometimes in an arbitrary way. Even though, this simplification may be justifiable at times, it is important to choose the model parameters in a systematic away and use an adequate input feature selection method. In this study, we present the development of a supervised learning model to predict the ground geology for the tunnel of Porto Metro Line C (Sousa and Einstein 2012), which includes a systematic way of choosing model input features. For this, we considered all the construction parameters recorded by the TBM as input features in the model to start, and then perform a systematic feature selection to reduce the presence of irrelevant features in the data and improve model accuracy and decrease learning times.

In addition, we have performed a simple sensitive analysis to investigate the influence of different TBM parameters in geology prediction. The performance metrics used to compare the different models in the sensitivity analysis were precision, recall and f1-score. Finally, the results are discussed to provide insights on how to perform model feature selection in ground forecast models for TBM construction.

2 Project Description

The data used in this study are from a tunnel built for the Porto metro line C (see Sousa and Einstein 2012). The project under study is a 2.3 km long from Campanhã to Trindade (Fig. 1). The tunnel was constructed using EPBM, a type of TBM which uses the pressurized excavated ground as means of face support. This EPBM could operate in open and closed modes in mixed ground conditions. The geology at the site of the Porto Metro tunnel is a highly weathered heterogenous granitic formation. The designer of the project specified six homogenous geomechanical groups (ground classes) based on the weathering degrees starting with intact granite as g1 to residual soils as g6. Based on the weathering degree, the geomechanical groups from g1 to g4 are rock like while g5 & g6 are soil like (see Table 1).

3 Data Description and Preparation

The data considered in this study ranges from ring 354 to ring 1611 (corresponding to station 0 + 656 until 0 + 2418 km, see Fig. 1), each ring is 1.4 m. 182 operation parameters were recorded by the EPBM, every 10 s. Figure 2 shows an example of a time series operation parameter of a given ring. To build our models, the time series data corresponding only to the time when the machine is advanced (i.e., advance rate is different than zero) was averaged for each ring and for each TBM parameter. The ground label of each ring was selected based on the location of the cutterhead after excavating the ring. The ground label is rock if the cutterhead is located in a rock like material (g1-g4), soils if the cutterhead is in soil like material (g5-g6) and mixed when the cutterhead is in a combination of soil and rock like materials. The total number of each ground class is shown in Fig. 3.

It is worth to mention, that the labeling in a EBPM tunnel has always a degree of uncertainty since the operator is unable to see the face. Face maps are done at certain locations of the tunnel and the remaining labels are determined based on interpolation of these face maps and observations by the operator (visual examination of the muck and by measurements of the weights of the muck). For this case study we had collected 40 face mappings. Figure 4 shows an example of a face mapping at Porto Metro.

Fig. 1. Porto Metro Line C Layout (adapted from Sousa and Einstein 2012)

Table 1. Porto Metro Line C geology formation geomechanical groups

Geo-mechanical group	Weathering Degree	GSI
G1	W1	68–85
G2	W2	45–65
G3	W3	30–45
G4	W4	15–30
G5	W5	<20
G6	W6	–

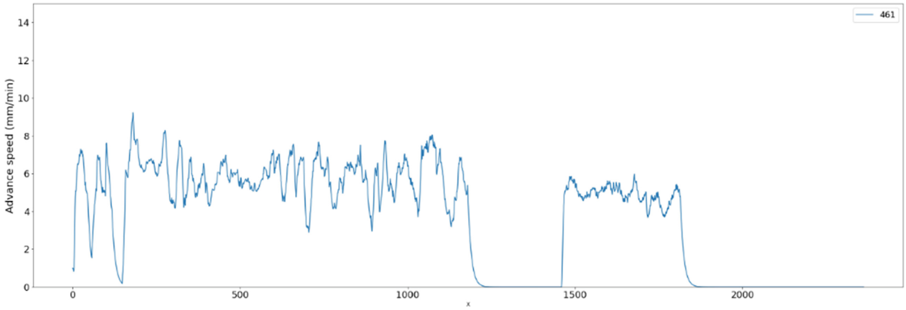

Fig. 2. Example of monitored advance speed time series for a ring. It includes both data from when the machine is advancing (advance speed nonzero) and idle (advance speed equal to zero)

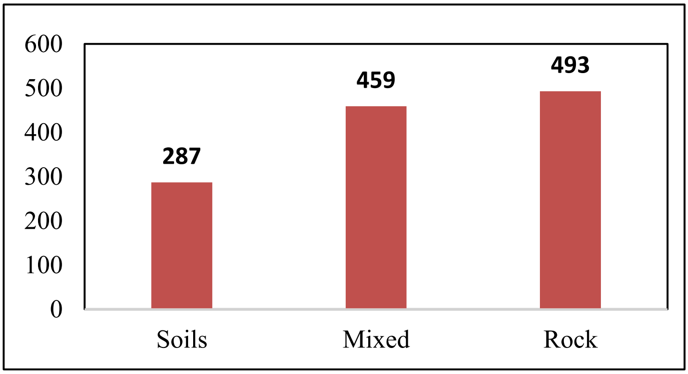

Fig. 3. Histogram of the available data labels

4 Feature Selection and Data Preprocessing

Feature selection and data preprocessing are critical steps in developing a machine learning model. The quality of the model depends mainly on the quality of the fed data. Feature selection, which consists of selecting the input features or variables of the

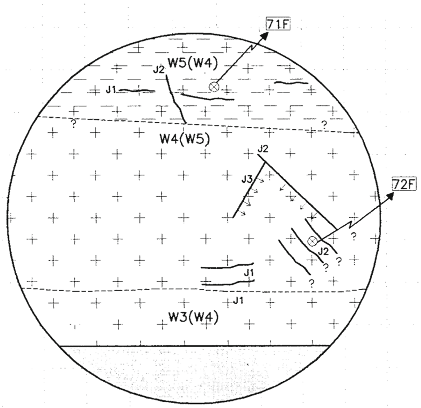

Fig. 4. Example of a face mapping at Porto Metro Line C (private correspondence). The maps contain information on the weathering degree of the exposed face. The weathering degree can be directly related to the geomechanical groups (see Table 1)

model, is an important technique in data pre-processing and is becoming an indispensable component of machine learning. However according to most existing studies there are no clear criteria for feature selection of ground prediction models. In this study, all the TBM parameters were initially considered for the model. Redundant parameters, i.e., highly correlated parameters, were removed with no loss of information. To do this, the first step consisted of plotting heatmaps for all parameters. Correlation heatmaps are used to show the correlation between pairs of variables. The correlation values are represented by a color scale (as shown in Fig. 5). For our case study, we looked at two cases: 1. Highly correlated parameters that are measured at different locations of the cutterhead (earth pressure, foam pressure and flow) and 2. Redundant features (parameters with high correlation but that were not the same parameter measured at different locations (e.g., advance rate and flow of excavated material). In both cases parameters were reduced to one. For example, for case 1, the earth pressure is measured by seven (7) sensors at different locations of the cutterhead. Since the heatmaps showed a high correlation between values measured at different sensors, only values from one sensor were kept for modelling purposes. For case 2, any parameters with correlation greater than 0.8 or smaller than −0.8 were removed and only one was selected to represent these parameters. Sixty-two (62) parameters remained to be included in the model after this process. Once the input features were selected, the data processing progressed as follows:

Cleaning the data: eighteen (18) data points (rings) with missing labels were removed from the dataset. They represented only about 1.5% of the size of the dataset and thus their removal does not to affect the model.
Removing outliers: outliers were removed for each parameter. The upper and lower bounds were Q1 - 1.5* IQR and Q3 + 1.5*IQR successively where.
Q1: Upper quartile
Q3: Lower Quartile
IQR: Interquartile, i.e., difference between the lower and the upper quartile.

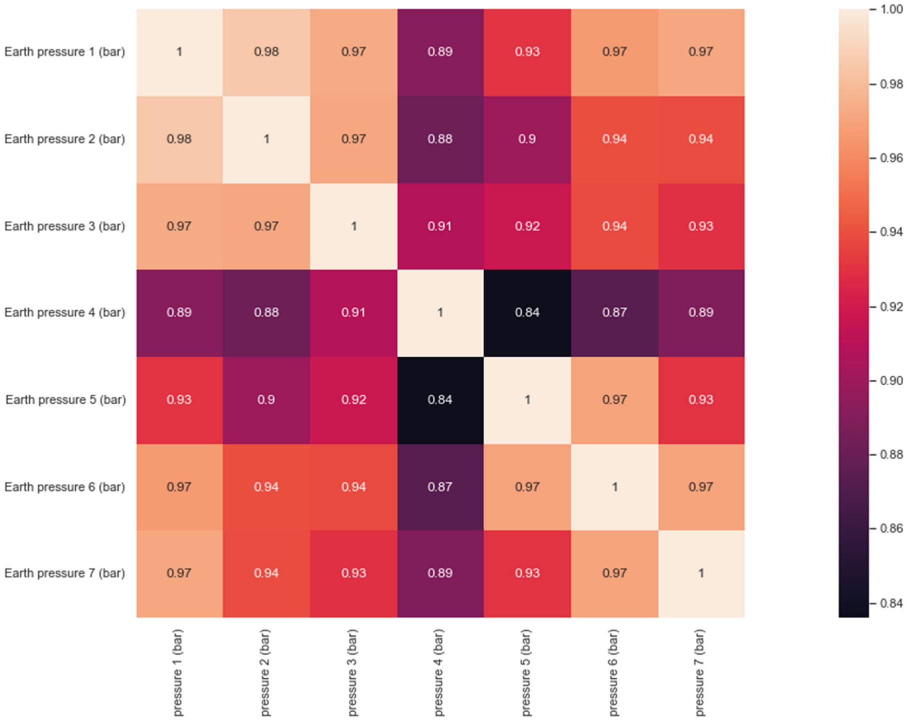

Fig. 5. Example of a Heatmap. Any correlation between two features above .80 was considered highly correlated and consequently one of the features was removed. The exception was made for the case where there were many sensors measuring the same parameter but at different locations of the cutterhead, like the map represented in the figure. In these cases, only the data of one sensor was kept.

The missing values were replaced by an average value of the previous and following rings.
Normalization of the data: all parameters were normalized to an interval between 0 and 1. Normalization is important and commonly used when the input data has variable scales as is the case of the TBM data.
Data Splitting: the available data were split into 80% training dataset and 20% as testing dataset. Both training and testing datasets have equal ground classes proportions.

5 Geology Prediction Model

XGBoost (Extreme Gradient Boosting) is one of the most popular and efficient algorithms used in supervised learning problems. It is an ensemble method based on the boosting technique (Chen and Guestrin 2016). The basic idea of boosting is to build multiple weak learners, normally decision trees, in series, to iteratively convert them into strong ones. Figure 6 is an illustration of how the XGBoost algorithm works.

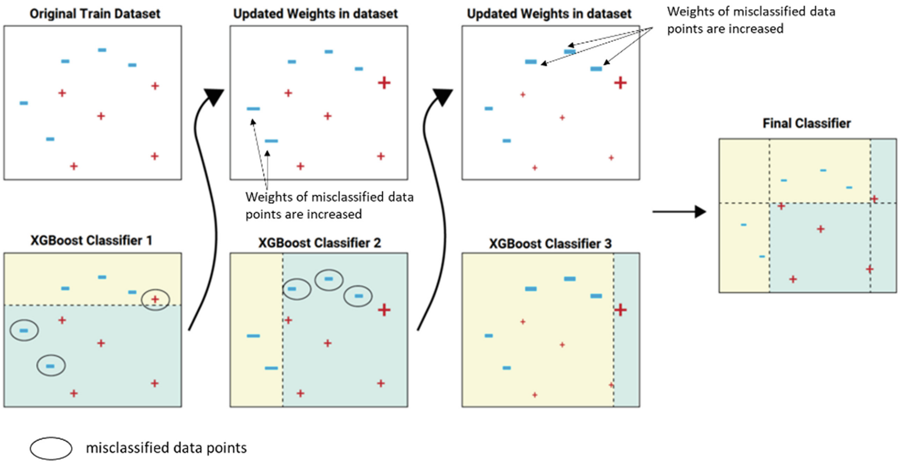

Fig. 6. Illustration of the principles of CGBoost (adapted from XG-Boost-FINAL-01 - Flirting with Models (thinknewfound.com)). The technique uses weak classifiers to iteratively construct a strong one. A weak learner or weak classifier is a one which performs just barely above chance. Initial weights are uniformly distributed in the original training set. After the predictions of the first weak classifier (in the figure XGBoost Classifier 1), more weight is assigned to the weight of the misclassified datapoints. This is done so that the next weak classifier, XGBoost Classifier 2, focuses its attention on these data points. Then after the predictions of the second weak classifier (in the figure XGBoost Classifier 2), more weight is assigned to the weight of the misclassified datapoints and so forth. The final classifier is a weighted combination of the weak classifiers.

XGBoost algorithm has various hyperparameters that need to be specified. Hyperparameters are the parameters that control the learning process, in contrast with other model parameters they cannot be inferred from data. For example, in neural networks, a hyperparameter is the size of the neural network. They don't affect the performance of the model, but they affect the speed and quality of the learning process, thus their importance.

Grid search was used to select the hyperparameters to optimize the learning model. Several values are selected for each hyperparameter, and Grid Search is used to run models for all different combinations of these hyperparameters. Finally, the hyperparameters of the model with the highest score are selected to be applied to the test dataset. Table 2 shows the values used for the hyperparameters of the XGboost model.

Four models were developed for this study using XGBoost. The first one was built using the sixty-two (62) TBM parameters obtained after the heatmaps study. Then feature importance was determined for all parameters, and features with importance less than 1% are eliminated. To calculate the feature importance, we used XGBoost's built-in function, which is computed by the gain, i.e., the relative contribution of a feature to in the model (for more details on the feature importance algorithm see Hastie et al. 2009).

The forty-two (42) parameters that remained after feature importance were then used to build a second model (model 2) to observe the effect of the eliminated features on model performance. The third and fourth models were developed after eliminating

Table 2. Hyperparameters Selected for XGBoost model

Hyperparameter	Value
Eval metric	mlogloss
n_estimators	75
Subsample	1
Colsample_bytree	0.4
Learning rate	0.01
Max depth	9

the ground conditioning parameters such as the foam, grout, and bentonite parameters to examine their effect on the model performance. The third (model 3) and fourth model (model 4) consisted of twenty-eight (28) and thirty-four (34) TBM parameters, respectively. Model 3 contained no foam parameters and Model 4 contained no grout or bentonite parameters. We wanted to verify if these conditioning parameters (that are set by the operator, not a reaction of the ground to the TBM) affected model performance.

Confusion matrices were plotted for the results of the models. A confusion matrix shows the numbers of correct vs incorrect predicted samples. The metrics used to assess the models are the precision, the recall, and the f1-score. The f1 score is used to combine both precision and recall in one metric. The formulas used to calculate each metric are as follows:

$$Precision = \frac{TP}{TP + FP} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$f1\ score = \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

where,
TP: True Positives
FP: False Positives
FN: False Negatives

6 Results and Discussion

The results of the models' performance are presented in Table 3 and Table 4, and the corresponding confusion matrices as shown in Figs. 7, 8, 9, and 10 which show the predicted label (ground class) versus the true label (ground class). Based on these matrices the prediction and recall presented in Table 4 were calculated using Eqs. 1 and 2. One can observe that Model 2 has the best performance overall and across all classes. Model

Table 3. Number of parameters and f1-score of different models

Metric/Ground Type	Model 1	Model 2	Model 3 (No Foam Parameters)	Model 4 (No Bentonite/Ground Parameters)
Number of Parameters	62	44	26	34
Weighted * f1-Score	0.901	0.925	0.865	0.898

*Weighted average of the f1-score of each ground class

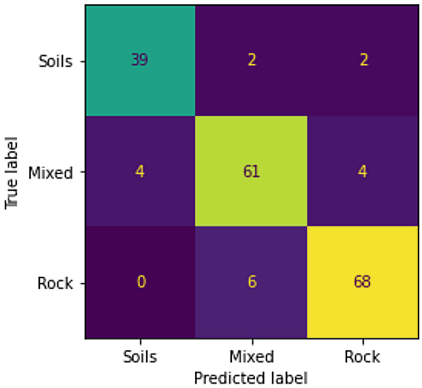

Fig. 7. Confusion Matrix of the Model 1.

3 (no foam parameters) performs worse in soil and particularly mixed, when compared with Model 1 and 2, indicating a strong relation between foam parameters and these two ground classes. Model 4 (no grout/bentonite parameters) performs slightly better in classifying Soil and Mixed than Model 3, but worse than Model 1 and 2, indicating again a strong relationship between the grout conditioning parameters and the ground classes. Overall, the predictions of ground class Rock seem to be less affected in the sensitivity analysis. This is confirmed by the precision values reported in Table 4. Table 3 shows the weighted f1-score for the four models to compare the overall performance of the models. The results show that the second model (i.e., the model obtained after removing all parameters with less than 1% importance) has the highest f1-score equals to 0.925 compared to 0.901 for the first model, which considers all parameters after heatmap selection. This confirms that feature importance ranking (FIR) is an extremely important step in supervised learning not only to improve efficiency and effectiveness of a predictive but also provide insight into the model, i.e. what features (variables) most influence the model's predictive power, and which features are irrelevant. Skipping this important step may result in a less robust and less accurate model, as well as slowing down the learning process.

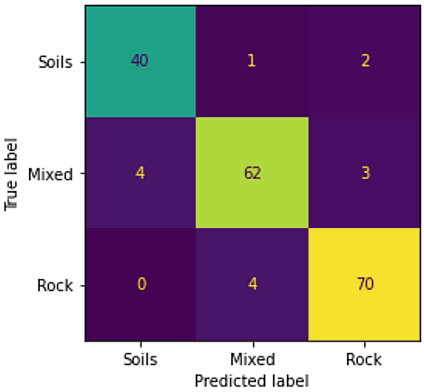

Fig. 8. Confusion Matrix of the Model 2. This model has the best performance across all ground classes.

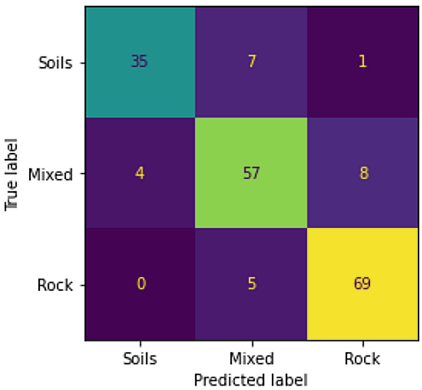

Fig. 9. Confusion Matrix of the Model 3.

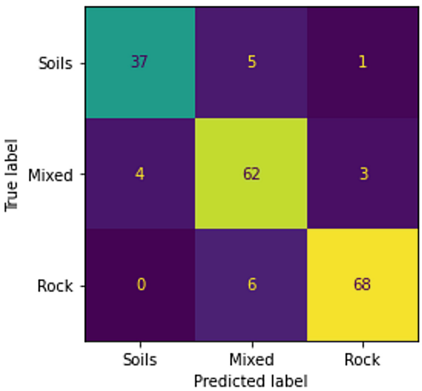

Fig. 10. Confusion Matrix of the Model 4.

Table 4. Recall and Precision of different models per ground class

Model/Ground Class	Soil		Mixed		Rock	
	Recall	Precision	Recall	Precision	Recall	Precision
Model 1	0.907	0.907	0.884	0.884	0.919	0.919
Model 2	0.930	0.910	0.900	0.925	0.946	0.943
Model 3	0.814	0.897	0.814	0.826	0.932	0.885
Model 4	0.860	0.900	0.900	0.850	0.918	0.940

7 Conclusions

In this paper, XGBoost model was used to develop machine learning models to predict the geology along the Porto metro tunnel line C using TBM parameters as input features. All the TBM parameters were initially considered for the analysis. Heatmaps were plotted for all the parameters to show linearly correlated parameters and eliminate them. Three other models were developed for this study. The first was by eliminating parameters with importance less than 1%. In the two other models, the ground conditioning parameters, foam parameters, and bentonite and grout parameters, were eliminated to examine their influence on the performance of the model. The performance of different models was compared using precision, recall and f1-score metrics. The following conclusions can be drawn from this study:

1. Model 2 had the best performance out of the four models with f1-score equal to 0.925. This shows that it is good practice to perform feature selection ranking (FIR), as including all TBM parameters may affect the model's performance negatively.
2. The rock ground class always has the highest scores for different models (see Figs. 7, 8, 9, and 10 and Table 4), and only gets mislabeled as Mixed. The Soil and Mixed classes have lower precision and get mislabeled both as Soil and Rock (in case of Mixed) and as Mixed and Rock (for the case of Soil). This seems to indicate that the machine parameter values in the case of Soil and Mixed are more variable. This led us to believe that other ground classes need to be refined or subclassified into more classes. This makes sense since the geology at Porto Metro is a granitic formation that has been weathered to different degrees. According to the results the rock like formation seems to possess more homogeneous properties, thus more consistent values are observed in the sensed data, and the mixed and soil like formations seem to be more heterogenous as a result of the weathering process, and more variable values are observed in the sensed data.
3. The results of Model 3 and Model 4 show that ground conditioning parameters play an important role in the classification of Soil and Mixed ground classes. For example, in the absence of foam parameters as input (Model 3), some Soil was misclassified as Mixed while some Mixed was classified as Rock. The ground conditioning parameters did not have a great effect on the prediction of the Rock.

References

Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2009)

Chan, M.H.C.: A Geological Prediction and Updating Model in Tunneling. M.S. thesis, Massachusetts Institute of Technology (1981). http://hdl.handle.net/1721.1/15886

Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. arXiv.org. 19 Aug 2022 (2016). https://arxiv.org/abs/1603.02754v2

Guan, Z., Deng, T., Du, S., Li, B., Jiang, Y., 2012. Markovian geology prediction approach and its application in mountain tunnels. Tunnelling and Underground Space Technology 31, 61–67. https://doi.org/10.1016/j.tust.2012.04.007

Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics (2009)

Hou, S., Liu, Y., Yang, Q., 2021. Real-time prediction of rock mass classification based on TBM operation big data and stacking technique of ensemble learning. Journal of Rock Mechanics and Geotechnical Engineering. https://doi.org/10.1016/j.jrmge.2021.05.004

Jung, J.H., Chung, H., Kwon, Y.S., Lee, I.M., 2019. An ANN to Predict Ground Condition ahead of Tunnel Face using TBM Operational Data. KSCE Journal of Civil Engineering 23, 3200–3206. https://doi.org/10.1007/s12205-019-1460-9

Liu, Q., Wang, X., Huang, X., Yin, X., 2020. Prediction model of rock mass class using classification and regression tree integrated AdaBoost algorithm based on TBM driving data. Tunnelling and Underground Space Technology 106. https://doi.org/10.1016/j.tust.2020.103595

Shi, M., Sun, W., Zhang, T., Liu, Y., Wang, S., Song, X.: Geology prediction based on operation data of TBM: comparison between deep neural network and soft computing methods; Geology prediction based on operation data of TBM: comparison between deep neural network and soft computing methods (2019)

Sousa, R.L., Einstein, H.H., 2021. Lessons from accidents during tunnel construction,Tunnelling and Underground Space Technology 113, 103916. https://doi.org/10.1016/j.tust.2021.103916.

Sousa, R.L., Einstein, H.H., 2012. Risk analysis during tunnel construction using Bayesian Networks: Porto Metro case study. Tunnelling and Underground Space Technology 27, 86–100. https://doi.org/10.1016/j.tust.2011.07.003

Wu, Z., Wei, R., Chu, Z., Liu, Q., 2021. Real-time rock mass condition prediction with TBM tunneling big data using a novel rock–machine mutual feedback perception method. Journal of Rock Mechanics and Geotechnical Engineering. https://doi.org/10.1016/j.jrmge.2021.07.012

Yin, X., Liu, Q., Huang, X., Pan, Y., 2022. Perception model of surrounding rock geological conditions based on TBM operational big data and combined unsupervised-supervised learning. Tunnelling and Underground Space Technology 120, 104285. https://doi.org/10.1016/j.tust.2021.104285

Zhang, Q., Liu, Z., Tan, J., 2019. Prediction of geological conditions for a tunnel boring machine using big operational data. Automation in Construction 100, 73–83. https://doi.org/10.1016/j.autcon.2018.12.022

Zhang, Q., Yang, K., Wang, L., Zhou, S., 2020. Geological Type Recognition by Machine Learning on In-Situ Data of EPB Tunnel Boring Machines. Mathematical Problems in Engineering 2020. https://doi.org/10.1155/2020/3057893

Zhao, J., Shi, M., Hu, G., Song, X., Zhang, C., Tao, D., Wu, W., 2019. A Data-Driven Framework for Tunnel Geological-Type Prediction Based on TBM Operating Data. IEEE Access 7, 66703–66713. https://doi.org/10.1109/ACCESS.2019.2917756